DOCUMENT RESUME

ED 171 744 TM 008 810

AUTHOR Mercer, Maryann
TITLE Answer Changing and Students' Test Scores.
PUB DATE Apr 79
NOTE 12p.; Paper presented at the Annual Meeting of the American Educational Research Association (63rd, San Francisco, California, April 8-12, 1979);

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Achievement Tests; Age Differences; *Answer Sheets; Content Analysis; Elementary Education; Low Ability Students; Mathematics; *Objective Tests; Reading; *Response Style (Tests); *Scores; Sex Differences
IDENTIFIERS *Answer Changing (Tests); Out of Level Testing

ABSTRACT

The effects of objective test answer changing on total test scores of students in grades 4 through 7, and the influences of sex, grade level, test content, and their interrelationships were examined. Standardized tests in reading and mathematics were administered at grade level and out-of-level. Answer changing was assessed by observing responses and erasures on the answer sheets. Results indicated: (1) answer changing increased total test scores; (2) high scoring students did not make significantly fewer answer changes than low scoring students; (3) males made significantly more wrong-right answer changes than females in mathematics tests at grade level; (4) sex and the interaction between sex and grade level did not influence the proportion of total answer changes; (5) sex, grade level, and their interaction did not influence the proportion of total wrong-right answer changes; and (6) test content appeared to be related to the proportion of total wrong-right answer changes, although there was no consistent pattern over the student/test grade level groups. (MH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED171744

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ANSWER CHANGING AND STUDENTS' TEST SCORES

Maryann Mercer
Rutgers University

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Maryann Mercer

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

TM 008 810

Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April, 1979.

ANSWER CHANGING AND STUDENTS' TEST SCORES

Maryann Mercer
Rutgers University

Researchers have investigated the effect of answer changing on objective test scores for over fifty years. Mueller and Wasser (1977) cite 17 studies which conclude that answer changing tends to improve students' test scores. Additional studies, not included in their review, have also substantiated this conclusion (Crocker & Benson, Note 1; Davis, 1975; Lynch & Smith, 1975; McMorris & Leonard, Note 2; Pascale, 1974; Smith & Moore, 1976). However, with the exception of the Crocker and Benson study, none of the research involved elementary or secondary school students, and the generalization of the findings to these younger populations was precluded.

The purpose of this study was to determine the effect of answer changing on the total test scores of elementary school students and to determine what differences, if any, exist due to the influences of sex, grade level, test content, and their interactions. Specifically, the following hypotheses were examined:

1. Elementary school students who change answers increase their total test scores.
2. High scoring students make fewer answer changes than low scoring students.
3. The proportion of W-R answer changes to total answer changes on verbal tests is higher for females than males and on mathematics tests is higher for males than females.

4. The proportion of total answer changes to total test items decreases over grade levels and is greater for males in higher grades than females.

5. The proportion of total W-R answer changes to total answer changes increases over grade levels and is greater for females in higher grades than males.

6. The proportion of W-R answer changes to total test items is related to test content.

Method

Sample

Random samples of students were selected from grades 4, 5, 6, and 7 in an urban midwestern public school system. Since students in this school system could be tested out-of-grade level, the selection was as follows: (1) 50 females and 50 males each from grades 4, 5, 6, and 7 who took the fourth grade level test and (2) 50 females and 50 males each from grades 5, 6, and 7 who took their respective grade level tests. The majority of students selected were one or more reading grade levels below their actual grade in school, attended schools in low SES communities, and represented minority racial/ethnic populations.

Procedure

Four subtests of the Iowa Tests of Basic Skills (reading, vocabulary, mathematics concepts, and mathematics problem solving) were administered. Answer changing was assessed by observing responses and erasures on students' answer sheets. For individual students, the number and type of answer changes were tallied, and their prechange and postchange subtest and total test scores (number right) were computed. For some hypotheses, individual groups of students were studied. For others, however, the student

groups were regrouped for separate analyses into the "4th Grade Group" (students who took fourth grade test) and the "At Level Group" (students who took grade level tests).

Since the number of test items varied by grade level and by subtest, proportions, normalized through the use of the arc sine transformation (Walker & Lev, 1953), were utilized in several analyses. The statistical procedures used depended on the particular hypothesis: specifically, Hypothesis 1 involved matched pairs *t* tests; Hypothesis 2, Wilcoxon rank sum tests; Hypothesis 3, *t* tests; and Hypotheses 4, 5, and 6, analyses of variance (6 was repeated measures).

Results

The findings can be summarized as follows:

1. The results, presented in Table 1, indicated that post-change scores were significantly higher than prechange scores (most had $p < .01$) for all subtests and the total test for each student/test grade level group. Thus, elementary school students who changed answers increased their total test scores.

2. There were no significant differences in the total number of answers changed by the 20% highest and 20% lowest scoring students identified in each student/test grade level group. High scoring students did not make significantly fewer answer changes than low scoring students.

3. For both the 4th Grade and At Level Groups, there were no significant differences between males and females on verbal tests. Females did not have a significantly higher proportion of W-R

answer changes than males on either the vocabulary or reading subtests. For mathematics tests, there were no significant differences between the sexes for the 4th Grade Group. For the At Level Group, there were significant differences ($p < .05$) between males and females on both subtests. (See Table 2.) Males did have a significantly higher proportion of W-R answer changes than females on the mathematics concepts and problem solving subtests when tested at grade level.

4. For the 4th Grade and At Level Groups, two way analyses of variance for sex, grade level, and their interaction were preformed. As indicated in Table 3, only the main effect of grade level was significant ($p < .01$) for the 4th Grade Group. Duncan's multiple range test revealed that the fifth grade students had a significantly higher ($p < .05$) mean proportion of total answer changes than students in the other grades. No other mean comparisons were significant. There were no significant main effects or interaction for the At Level Group. The test for a linear trend over grade levels was significant ($p < .05$) for the 4th Grade Group, but not the At LevelGroup. Sex and the interaction of sex and grade level were not factors influencing the proportion of total answer changes. Grade level appeared to be a factor for students who took the fourth grade level test.

5. For the 4th Grade and At Level Groups, the results of of the two way analyses of variance for sex, grade level, and their interaction indicated that there were no significant main effects or interaction for either group. The test for a linear

trend over grade levels was not significant for either group. In this study, sex, grade level, and their interaction were not factors influencing the proportion of total W-R answer changes.

6. As Table 4 indicates, several student/test level groups had significant differences ($p < .05$ or more) among proportions of W-R answer changes for the subtests. Duncan's multiple range tests revealed that for fifth and sixth grade students who took the fourth grade test and for sixth and seventh grade students who were tested at grade level, one or more of the subtests had significantly higher ($p < .05$) mean proportions of W-R answer changes than the other subtests. In three cases, the reading subtests had the highest proportion. No other mean comparisons were significant. Test content did appear to be related to the proportion of total W-R answer changes, although there was no consistent pattern over the student/test grade level groups.

Discussion

Although this study did find significant differences when test content was considered, a consistent pattern was not established. Additional investigation will be necessary to determine if factors subsumed under the rubric of test content cause the differences. Similarly, the inconclusive results related to sex and grade level differences suggest that there are other moderating variables which should be considered. Also, at the present time, students are not encouraged to change answers. If they were, the gains observed might not remain.

This study has provided additional empirical data on answer

changing which has confirmed some previous findings, explored new issues, and suggested further avenues of research. The results of the study have implications for classroom teachers and other test administrators because answer changing tends to raise test scores. Advising students not to change answers may cause needlessly lower test scores. No advice would be a wiser course of action.

Reference Notes

1. Crocker, L., & Benson, J. Effects of examinee response changes on item and test characteristics. Paper presented at the annual meeting of the American Educational Research Association, New York, 1977.

2. McMorris, R. F., & Leonard, G. Item responses changes and cognitive style. Paper presented at the annual meeting of National Council on Measurement in Education, San Francisco, 1976.

References

Davis, R. E. Changing examination answers: An educational myth? Journal of Medical Education, 1975, 50, 685-687.

Lynch, D. O., & Smith, B. C. Item response changes: Effects on test scores. Measurement and Evaluation in Guidance, 1975, 7, 220-224.

Mueller, D. J., & Wasser, V. Implications of changing answers on objective test items. Journal of Educational Measurement, 1977, 14, 9-13.

Pascale, P. J. Changing initial answers on multiple-choice achievement tests. Measurement and Evaluation in Guidance, 1974, 6, 236-238.

Smith, A., & Moore, J. C. The effects of changing answers on scores of non-test-sophisticated examinees. Measurement and Evaluation in Guidance, 1976, 8, 252-254.

Walker, H. B., & Lev, J. Statistical inference. New York: Holt, Rinehart, & Winston, 1953.

Table 1

Mean Difference Scores (Pre vs. Postchange) for Subtests and Total Test

Student/Test Grade Level[a]		M	SD	t[b]
4/4	Vocabulary	.49	1.08	4.55
	Reading	1.11	1.94	5.71
	Math Concepts	.57	1.28	4.45
	Math Problem	.42	.87	4.85
	Total	2.59	2.90	8.92
5/4	Vocabulary	.57	1.55	3.69
	Reading	1.88	2.68	7.02
	Math Concepts	.85	1.59	5.36
	Math Problem	.43	.99	4.36
	Total	3.73	3.67	10.16
6/4	Vocabulary	.30	.99	3.03
	Reading	1.10	1.99	5.52
	Math Concepts	.51	1.22	4.19
	Math Problem	.37	.72	5.14
	Total	2.28	2.72	8.40
7/4	Vocabulary	.31	1.03	3.00
	Reading	.81	1.52	5.34
	Math Concepts	.31	.77	4.00
	Math Problem	.38	.91	4.19
	Total	1.81	2.46	7.37
5/5	Vocabulary	.62	1.70	3.64
	Reading	1.15	2.27	5.07
	Math Concepts	.61	1.00	6.08
	Math Problem	.28	.92	3.04
	Total	2.66	3.46	7.68
6/6	Vocabulary	.65	1.59	4.10
	Reading	1.19	1.63	7.30
	Math Concepts	.58	1.22	4.74[c]
	Math Problem	.21	.89	2.36[c]
	Total	2.63	2.95	8.92
7/7	Vocabulary	.44	1.19	3.96
	Reading	1.14	1.87	6.10
	Math Concepts	.79	1.39	5.69
	Math Problem	.31	.86	3.60
	Total	2.68	3.66	7.32

[a] 100 students per student/test grade level.

[b] $p < .01$ unless otherwise specified.

[c] $p < .05$.

Table 2

Results of t Test on W-R Answer Changes of Males and Females for Mathematics Concepts and Problem Solving Subtests

Group	n	M	SD	t
4th Grade				
Concepts				
Males	200	1.08	1.34	-.06
Females	200	1.09	1.26	
Problem Solving				
Males	200	1.05	1.38	.81
Females	200	.94	1.32	
At Level				
Concepts				
Males	200	1.32	1.32	1.71*
Females	200	1.10	1.27	
Problem Solving				
Males	200	.97	1.32	1.66*
Females	200	.75	1.26	

*p <.05.

Table 3

ANOVA Results for Total Answer Changes

Source	df	SS	MS	F
4th Grade				
Sex	1	.004	.004	.21
Grade	3	.42	.14	7.51*
Sex X Grade	3	.03	.009	.45
Error	392	7.31	.02	
At Level				
Sex	1	.05	.05	2.83
Grade	3	.10	.03	1.82
Sex X Grade	3	.002	.0008	.04
Error	392	7.20	.02	

*p < .01.

Table 4

ANOVA Results for W-R Answer Changes on Subtests by Students (Repeated Measures)

Student/Test Grade Level[a]		df	SS	MS	F
4/4	Subtest	3	.32	.11	2.23
	Student	99	5.47	.06	1.16
	Error	297	14.16	.05	
5/4	Subtest	3	1.01	.34	6.34**
	Student	99	7.71	.08	1.46
	Error	297	15.86	.05	
6/4	Subtest	3	.36	.12	2.99*
	Student	99	5.71	.06	1.45
	Error	297	11.85	.04	
7/4	Subtest	3	.12	.04	1.13
	Student	99	5.25	.05	1.45
	Error	297	10.87	.04	
5/5	Subtest	3	.17	.06	1.58
	Student	99	6.88	.07	1.94
	Error	297	10.62	.04	
6/6	Subtest	3	.83	.28	7.43**
	Student	99	6.21	.06	1.69
	Error	297	11.03	.04	
7/7	Subtest	3	.58	.19	6.17**
	Student	99	6.13	.06	1.98
	Error	297	9.30	.03	

[a]100 students per student/test grade level.

*$p < .05$.

**$p < .01$.